ABSTRACT OF THE DISCLOSURE

Disclosed is an apparatus for and method of extracting or mining data from a visually displayable file using quads to enclose textual symbols and graphics where tables of data are detected and processed differently than straight text or text and graphics. Quads enclosing

5   textual symbols and graphics are oversized and overlapping quads are assembled into frames. Each frame is then treated as a separate title, paragraph, graphic or table in accordance with an analysis of the extracted material. An output (mined) document is then created that may be used by word processing and web browser programs to display, print or transport the mined material.